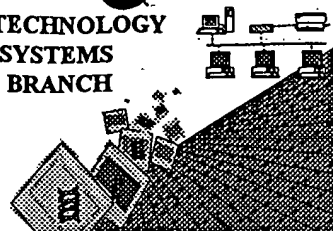


BIOTECHNOLOGY
SYSTEMS
BRANCH



RAW SEQUENCE LISTING
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/050,000
Source: PCT 10
Date Processed by STIC: 03/05/02

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER**
VERSION 3.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND
TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. Hand Carry directly to:
U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name,
Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
Or
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two,
2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office,
Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 01/29/2002

Raw Sequence Listing Error Summary

PCT 10

ERROR DETECTED

SUGGESTED CORRECTION

SERIAL NUMBER: 10/050,000

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1 Wrapped Nucleics
 Wrapped Aminos The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2 Invalid Line Length The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3 Misaligned Amino
 Numbering The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4 Non-ASCII The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5 Variable Length Sequence(s) contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6 PatentIn 2.0
 "bug" A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7 Skipped Sequences
 (OLD RULES) Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence:
 (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
 (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
 (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
 This sequence is intentionally skipped

 Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8 Skipped Sequences
 (NEW RULES) Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence.
 <210> sequence id number
 <400> sequence id number
 000
- 9 ✓ Use of n's or Xaa's
 (NEW RULES) Use of n's and/or Xaa's have been detected in the Sequence Listing.
 Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
 In <220> to <223> section, please explain location of n or Xaa; and which residue n or Xaa represents.
- 10 Invalid <213>
 Response Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
- 11 Use of <220> Sequence(s) missing the <220> "Feature" and associated numeric identifiers and responses.
 Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
 (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12 PatentIn 2.0
 "bug" Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13 Misuse of n n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.



Does Not Comply
Corrected Diskette Needed

PCT10

Error on p. 2

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/050,000

DATE: 03/05/2002

TIME: 14:01:24

Input Set : A:\0365-0529P.ST25.txt

Output Set: N:\CRF3\03052002\J050000.raw

```

3 <110> APPLICANT: NAKARI-SETALA, Tiina
5 <120> TITLE OF INVENTION: A METHOD FOR DECREASING THE FOAM FORMATION DURING
CULTIVATION OF A
6 MICROORGANISM
8 <130> FILE REFERENCE: 0365-0529P
C--> 10 <140> CURRENT APPLICATION NUMBER: US/10/050,000
11 <141> CURRENT FILING DATE: 2002-02-20
13 <160> NUMBER OF SEQ ID NOS: 8
15 <170> SOFTWARE: PatentIn Ver. 2.1
17 <210> SEQ ID NO: 1
18 <211> LENGTH: 2868
19 <212> TYPE: DNA
20 <213> ORGANISM: Trichoderma reesei
22 <220> FEATURE:
23 <221> NAME/KEY: gene
24 <222> LOCATION: (1523)..(1950)
25 <223> OTHER INFORMATION: hfb1
27 <400> SEQUENCE: 1
28 ttgtatggc tggatctcga aaggcccttg tcatcgccaa gcgtggctaa tatcgaatga 60
29 gggacaccga gttgcatatc toctgatcat tcaaacgaca agtgtgaggt aggcaatcct 120
30 cgtatcccat tgctgggctg aaagcttcac acgtatcgca taagcgtctc caaccagtgc 180
31 ttaggtgacc cttaaggata cttacagtaa gactgtatta agtcagtcac tctttcactc 240
32 gggctttgaa tacgatcctc aatactcccg ataacagtaa gaggatgata cagcctgcag 300
33 ttggcaaatg taagcgtaat taaactcagc tgaacggccc ttgttgaaag tctctctcga 360
34 tcaaagcaaa gctatccaca gacaagggtt aagcaggctc actcttccta cgccttggat 420
35 atgcagcttg gccagcatcg cgcattggcca atgatgcacc cttcacggcc caacggatct 480
36 cccgttaaac tccoctgtaa cttggcatca ctcatctgtg atcccaacag actgagttgg 540
37 gggctgcggc tggcggatgt cggagcaaag gatcacttca agagcccaga tccggttgg 600
38 ccattgccaa tggatctaga ttcggcacct tgatctcgat cactgagaca tggtagattg 660
39 cccggaacga ccacaactcc cctgtgtcga ttgagtcccc atatgcgtct tctcagcgtg 720
40 caactctgag acggattagt cctcacgatg aaattaactt ccagcttaag ttcgtagcct 780
41 tgaatgagtg aagaaatttc aaaaacaaac tgagtagagg tcttgagcag ctgggggtgg 840
42 acgcccctcc togactcttg ggacatcgta cggcagagaa tcaacggatt cacacctttg 900
43 ggtcgagatg agctgatctc gacagatacg tgcttcacca cagctgcagc tacctttgcc 960
44 caaccattgc gttccaggat cttgatctac atcacgcgag cacccgagcc aggacggaga 1020
45 gaacaatccg gccacagagc agcaccgcct tccaactctg ctccctggcaa cgtcacacaa 1080
46 cctgatatta gatatccacc tgggtgattg ccattgcaga gaggtggcag ttggtgatac 1140
47 cgactggcca tgcaagacgc ggccgggcta gctgaaatgt ccccgagagg acaattggga 1200
48 gcgtctatga cggcgtggag acgacgggaa aggactcagc cgtcatgttg tgttgccaat 1260
49 ttgagattgt tgaccgggaa aggggggacg aagaggatgg ctgggtgagg tggattggg 1320
50 aggatgcata attgactcga gtgagcgtat tagagctcca agaataaaa tatcccttct 1380
51 ctgtcttctc aaaaatctct tccatcttgt ccttcacag caccagagcc agcctgaaca 1440
52 cctccagtc aacttccctta ccagtacatc tgaatcaaca tccattcttt gaaatctcac 1500
53 cacaaccacc atcttcttca aaatgaagtt cttcgccatc gccgtctctt ttgccgccgc 1560

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/050,000

DATE: 03/05/2002

TIME: 14:01:24

Input Set : A:\0365-0529P.ST25.txt

Output Set: N:\CRF3\03052002\J050000.raw

```

54 tgcctgttgc cagcctctcg aggaccgcag caacggcaac ggcaatgttt gccctcccgg 1620
55 cctcttcagc aacccccagt gctgtgccac ccaagtcctt gccctcatcg gccttgactg 1680
56 caaagtccgt aagttgagcc ataacataag aatcctcttg acggaaatat gccttctcac 1740
57 tcctttaccc ctgaacagcc tcccagaacg ttacgacgg caccgacttc cgcaacgtct 1800
58 gcgcaaaaac cggcgcccag cctctctgct gcgtggcccc cgttgtaagt tgatgcccc 1860
59 gctcaagctc cagtcttttg caaacccatt ctgacaccca gactgcaggc cggccagggt 1920
60 cttctgtgcc agaccgccgt cgggtgcttga gatgcccgcc cggggtcaag gtgtgcccg 1980
61 gagaaagccc acaaagtgtt gatgaggacc atttccggta ctgggaaagt tggctccacg 2040
62 tgtttgggca ggtttgggca agttgtgtag atattccatt cgtacgcat tcttattctc 2100
63 caatatttca gtacactttt cttcataaat caaaaagact gctattctct ttgtgacatg 2160
64 ccggaaggga acaattgctc ttggtctctg ttatttgcaa gtaggagtgg gagattcgcc 2220
65 tttagaagaa tagagaagct gtgcttgacc gtggtgtgac tcgacgagga tggactgaga 2280
66 gtgttaggat taggtcgaaac gttgaagtgt atacaggatc gtctggcaac ccacggatcc 2340
67 tatgacttga tgcaatggtg aagatgaatg acagtgtgaa aggaaaagga aatgtccgcc 2400
68 ttcaagctgat atccacgcca atgatacagc gatatacctc caatatctgt gggaacgaga 2460
69 catgacatat ttgtgggaac aacttcaaac agcgagccaa gacctcaata tgcacatcca 2520
70 aagccaaaca ttggcaagac gagagacagt cacattgtcg tcgaaagatg gcacgtacc 2580
71 caaatcatca gctctcatta tcgcctaaac cacagattgt ttgccgtccc ccaactccaa 2640
72 aacgttacta caaaagacat gggcgaatgc aaagacctga aagcaaacc tttttgcgac 2700
73 tcaattccct cctttgtcct cggaatgatg atccttcacc aagtaaaaga aaaagaagat 2760
74 tgagataata catgaaaagc acaacggaaa cgaaagaacc aggaaaagaa taaatctatc 2820
75 acgcaccttg tccccacact aaaagcaaca gggggggtaa aatgaaat 2868

```

78 <210> SEQ ID NO: 2

79 <211> LENGTH: 3585

80 <212> TYPE: DNA

81 <213> ORGANISM: Trichoderma reesei

83 <220> FEATURE:

84 <221> NAME/KEY: gene

85 <222> LOCATION: (1191)..(1593)

86 <223> OTHER INFORMATION: hfb2

88 <400> SEQUENCE: 2

```

89 ctgcagcagc tgaagcttgc atgcctgcac cctttgtgag cgactgcac cttttgcac 60
90 acactgccgt cgacgtctct cttccgacct tggccagctg gacaagcaac acaccaatga 120
91 cgctttgtat tattagagta tatgcaagtc tcaggactat cgactcaact ctaccacccg 180
92 aggacgatcg cggcacgata cgccctcggt ctcattggcc caagcagacc aactgcccct 240
93 ggagcaagat tcagcccaag ggagatggac ggcagggcac gccaggcccc caccaccaag 300
94 ccaactccctt tggccaaatc agcttgcacg tcaagagaca tcgagctgtg ccttgaaatt 360
95 actaacaacc agggatggga aacgaagcct gcttttgga agacaacaat gagagagaga 420
96 gagagagggg gagagacaat gagtgccaca aacctggtag tgctccgcca atgcgtctga 480
97 aatgtcacat ccgagtcctt gggcctctgt gagaatgtcc agagtaatac gtgttttgcg 540
98 aatagtcctc tttcttgagg actggatacc tacgataccc tttttgagtt gatgcggtgc 600
99 tttcgaagta ttatctggag gatagaagac gtctaggtaa ctacacaaaa ggcctatact 660
100 ttggggagta gcccaacgaa aggtaactcc tacggcctct tagagccgtc atagatccta 720
101 cagcctcttg gagccgtcat agatcacatc tgtgtagacc gacattctat gaataatcat 780
102 ctcatcatgg ccacatacta ctacatacgt gtctctgctt acctgacatg tagcagtggc 840
103 caagacacca aggccccagc atcaagcctc cctacctatc ccttccattg tacagcggca 900
104 gagagattgc gatgagccct ctccctacct acagacggct gacaatgtcc gtataccacc 960
105 agccaacgtg atgaaaacaa ggacatgagg aacagcctgc gagagctgga agatgaagag 1020
106 ggccagaaaa aaaagtataa agaagacctc gattcccgcc atccaacaat cttttccatc 1080

```

*must give locations of "n" and
 explain what residue "n"
 represents - see error
 summary sheet item 9*

RAW SEQUENCE LISTING

DATE: 03/05/2002

PATENT APPLICATION: US/10/050,000

TIME: 14:01:24

Input Set : A:\0365-0529P.ST25.txt

Output Set: N:\CRF3\03052002\J050000.raw

```

107 ctcatcagca cactcatcta caaccatcac cacattcact caactcctct ttctcaactc 1140
108 tccaaacaca aacattcttt gttgaatacc aaccatcacc acctttcaag atgcagttct 1200
109 tcgccgtcgc cctcttcgcc accagcgccc tggctgctgt ctgccctacc ggcctcttct 1260
110 ccaaccctct gtgctgtgcc accaacgtcc tcgacctcat tggcgttgac tgcaagaccc 1320
111 gtatgttgaa ttccaatctc tgggcatact gacattggac gatacagttg acttacacga 1380
112 tgctttacag ctaccatcgc cgtcgacact ggcgccatct tccaggctca ctgtgccagc 1440
113 aagggctcca agcctctttg ctgcgttgct cccgtggtaa gtagtgctcg caatggcaaa 1500
114 gaagtaaaaa gacatttggt cctgggatcg ctaactcttg atatcaaggc cgaccaggct 1560
115 ctctgtgcc agaaggccat cggcaccttc taaagcaatg gcttgcttta ctgccggcag 1620
116 tctttgagaa ctctgggctc acaaaagacg acttgcatgt atcatggggg ctcgcaaagt 1680
117 ggaggatttg gaggggattg aggctgggtt tggcctatta gaggattgca taatggaaga 1740
118 tttgcgagca ggacatagac gtatctagag ttctagtcaa tacattattg aaaagttgga 1800
119 gtatacctat cgctggcact ggtatcttga agatatcttc tcttcttggt aggttatgta 1860
W--> 120 tggcaatcag tcgaaatcta tttgaagaca gagctcaagc ttcaaacatt cacctgngaa 1920
121 ttgaccatth tgtttcgatg gttgcagttg gtgggtgtca cttctgcaat catgtacgag 1980
122 cacaagtata gcagtattcc atctgatctg catctgggta aatgtcgcca ctctacctag 2040
123 gtaccaata aataccgaat tggtcagctc tcgggtgaca aaccggcccg cttttcgacc 2100
W--> 124 gtgctctgtc caattctagg cttgtcaatg gttcctgact gtgataaacc ttggagctan 2160
125 cataacttac cttacaataa atccaaactgc cggcaacttg ttcctttcac ccaaccactc 2220
126 gcaaacatca cgcaacctgt ctcgatcccc tgtccgaaat ctgcttgcca acgtatcatc 2280
127 acaaatcata cacacagaca aaaaggagcc aaagcagcaa tggcaagaca ccgaggccgg 2340
128 cagcgcgccc gtgcgcggtt ttaaaaagcg aagcgcaag ggcaaagcca acctgcgcaa 2400
129 acgaacaacg aagccttccc ccgcgcgca ggcagacgca cagcgacagc gacttttctt 2460
130 cgtcggaaaga cgaagccggg cacagagtca agaggcgcaa gaggacggcc gtcgtcaccg 2520
131 ccgccgcgga ggggcgcgcg gccagcaac cgggacgacg gcggcggcgc aacagccgcc 2580
132 ttcacggcca acagaagcgt cccgattgct gacagcaacg acgcgacca gacagcaac 2640
133 tggtagcagc aggacgcaaa ggacgcgctc tcggcaaaga acctcctcgg atcttcgaga 2700
134 gcgtagcagg acgcgcagcc agacggcagc tacaagggcc tggcgaacca gacgtccttt 2760
135 atacaaaaga atccggatgc gcccgggaag acagttgggc ccgtcaaggc gcctaccaac 2820
136 atccgcaccg tcaccattac agattatgcc ccggacacgt gtaaagagtg agtttgcatc 2880
137 aatagccaga atccccccc ccgataccgt acattgagca tatgtgact cgtcataatc 2940
138 tttctagtta tcgcataacc ggctatataa gtactccct tttccatgat tattccagtc 3000
139 gcgtagtgac atttctagga gcctttactg tggttttggc gacaattgca agtatcttca 3060
140 cgcgagagaa gacctcaagg caggctggca gctggatcaa gagtgggaaa aggtcaccaa 3120
141 gggcaagaag aacctggggg gaacggtagt ggcagcgcg aaccggaaca aggccagggt 3180
142 ggacgagggc gacgacgacg acgacgaaga ggcgatgctc gagaacattc cgtttgcttg 3240
143 catcatctgc agggaatcgt acaaggagcc gattgtgacg aggtgcgggc actacttttg 3300
144 cctgccgtgc gctctgcagc ggtacaagaa ggatccgacg tgtgcggcgt gtggctcggg 3360
145 cacgaatggc gtgtttaatt cggcgacgag gttgaagaag ctgctggaga agaagaggga 3420
146 gagggcgggc aggaggagac aggaggcgat agagaggggc gaggaagtca gtgatgaaga 3480
W--> 147 ggaggaggag gaggaggact gatgatgatg gggcnagatg acgatgcagg tcgactctag 3540
148 agatcccccg taccgagctc gaattcatcg atgatatcag atccc 3585
151 <210> SEQ ID NO: 3
152 <211> LENGTH: 63
153 <212> TYPE: DNA
154 <213> ORGANISM: Artificial Sequence
156 <220> FEATURE:
157 <223> OTHER INFORMATION: Description of Artificial Sequence: 5' primer
159 <400> SEQUENCE: 3

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/050,000

DATE: 03/05/2002

TIME: 14:01:24

Input Set : A:\0365-0529P.ST25.txt

Output Set: N:\CRF3\03052002\J050000.raw

```

160 actacacgga ggagctcgac gacttcgagc agcccgagct gcacgcagag caacggcaac 60
161 ggc 63
164 <210> SEQ ID NO: 4
165 <211> LENGTH: 30
166 <212> TYPE: DNA
167 <213> ORGANISM: Artificial Sequence
169 <220> FEATURE:
170 <223> OTHER INFORMATION: Description of Artificial Sequence: 3'primer
172 <400> SEQUENCE: 4
173 tcgtacggat cctcaagcac cgacggcggg 30
176 <210> SEQ ID NO: 5
177 <211> LENGTH: 13
178 <212> TYPE: PRT
179 <213> ORGANISM: Artificial Sequence
181 <220> FEATURE:
182 <223> OTHER INFORMATION: Description of Artificial Sequence: linker
184 <400> SEQUENCE: 5
185 Pro Gly Ala Ser Thr Ser Thr Gly Met Gly Pro Gly Gly
186 1 5 10
189 <210> SEQ ID NO: 6
190 <211> LENGTH: 14
191 <212> TYPE: PRT
192 <213> ORGANISM: Artificial Sequence
194 <220> FEATURE:
195 <223> OTHER INFORMATION: Description of Artificial Sequence: linker
197 <400> SEQUENCE: 6
198 Gly Thr Leu Val Pro Arg Gly Pro Ala Glu Val Asn Leu Val
199 1 5 10
202 <210> SEQ ID NO: 7
203 <211> LENGTH: 48
204 <212> TYPE: DNA
205 <213> ORGANISM: Artificial Sequence
207 <220> FEATURE:
208 <223> OTHER INFORMATION: Description of Artificial Sequence: 5'primer
210 <400> SEQUENCE: 7
211 gaattcggta ccctcgctccc tcgcggtccc gccgaagtga acctggtg 48
214 <210> SEQ ID NO: 8
215 <211> LENGTH: 34
216 <212> TYPE: DNA
217 <213> ORGANISM: Artificial Sequence
219 <220> FEATURE:
220 <223> OTHER INFORMATION: Description of Artificial Sequence: 3'primer
222 <400> SEQUENCE: 8
223 tgaattccat atgctaacc cgtttcatct ccag 34

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/050,000

DATE: 03/05/2002

TIME: 14:01:25

Input Set : A:\0365-0529P.ST25.txt

Output Set: N:\CRF3\03052002\J050000.raw

L:10 M:270 C: Current Application Number differs, Replaced Current Application Number

L:120 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2

L:124 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2

L:147 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2